Address Validator Transformation Overview

The Address Validator transformation compares input address data with address reference data to determine the accuracy of input addresses and fix errors in those addresses. The transformation can create columns with new address data, corrected address data, and status information on each address.

By default, the Address Validator transformation is a passive transformation, but setting the value of **Execution Instances** property higher than one makes the transformation active.

You can validate and update address data in the following ways:

- Compare input addresses to address reference data provided by Informatica.

- Generate detailed status reports on the validity of each input address, its deliverable status, and the nature of any errors or ambiguities it contains.

- Fix errors and complete partial address records using information from the reference data. To fix an address, the transformation must find a positive match with an address in the reference data.

- Write output addresses in an envelope-ready format. You define the format when you select the output ports on the transformation.

- Add information that does not appear in the standard address but that helps postal delivery, such as geocoding information.

Address Reference Data

An address reference data set contains data that describes all deliverable addresses in a country. The address validation process searches the reference data set for the address that most closely resembles the input address data. If the process finds a close match in the reference data set, it writes new values for any incorrect or incomplete data values. The process creates a set of alphanumeric codes that describe the type of match found between the input address and the reference addresses.

Address validation can also restructure the address, and it can add information that is absent from the input address, such as a four-digit ZIP code suffix for a United States address.

Address reference data sets do not form part of the Data Quality installation. You must download the address reference data sets separately and install them using the Data Quality Content Installer. Use the Preferences window in the Developer tool to view information about the address reference data sets on your system.

Port Groups and Port Selection

The Address Validator transformation contains predefined port groups that contain the input and output ports you can use. When you configure an Address Validator transformation, you browse the groups and select the ports that you need.

Select input ports that correspond to the structure of the address input data. Select output ports that contain the address data that your project requires.

You can add input and output ports directly to the transformation, or you can create a default model that contains input and output ports. When you add ports directly to the transformation, the ports you select apply to that transformation only. When you add ports to the default model, the ports you select apply to future Address Validator transformations that you create.

Note: The default model is not a repository object. The default model resides on the machine you use to create it.

You can also add pass-through ports to the transformation for columns that you do not want the Address Validator transformation to process.

Address Validator Transformation Input Port Groups

Before you can connect address data to input ports on the transformation, you browse the input groups and select the ports that correspond to the structure and content of the input data.

Browse the output groups and select the ports that match your data requirements.

The transformation has the following input port groups:

- **Discrete.** A discrete port accepts a single address element. Connect address fields to these ports if each field represents a unique address element, such as house number, apartment number, street name, city, state, or ZIP code.

- **Multiline.** A multiline port accepts a string containing multiple address elements. Connect address fields to these ports if each field represents multiple address elements, for example when Address Line 1 contains "100 Main Street" and Address Line 2 contains "New York, NY 10001."

- **Hybrid.** A hybrid port accepts a string containing single or multiple address elements. Connect address fields to these ports if the address contains a mix of discrete and multi-line elements. A common format combines house number, street name, and apartment information in a single field and uses discrete fields for city, state, and ZIP code.

Address Validator Transformation Output Port Groups

Before you can connect the Address Validator transformation to other transformations or data objects, you determine the types of information you need and the structure that the output addresses will take.

Browse the output groups and select the ports that match your data requirements.

Note: You can select ports from multiple output groups, and you can select ports that have common functionality.

The transformation has the following predefined output groups:

- **Address Elements.** Use to write street address elements, such as house number, apartment number, and street name, to separate ports.

- **Last Line Elements.** Use to write locality information, such as post codes and city and state names, to separate ports.

- **Geocoding.** Use to write geo-coding information, such as latitude and longitude coordinates.

- **US Specific.** Use to write additional data on United States addresses that can help postal delivery, such as DPV data.

- **Canada Specific.** Use to write SERP report data for Canadian addresses.

- **Australia Specific.** Use to write AMAS report data for Australian addresses.

- **UK Supplementary.** Use to write Delivery Point Suffix data for United Kingdom addresses.

- **US Supplementary.** Use to write supplemental data such as county FIPS and state FIPS codes for United States addresses.

- **Country.** Use to write the country name or ISO country code.

- **Status Info.** Use to write information on the quality of each input address.

- **Formatted Address Line.** Use to write addresses that are formatted for mailing. Formatted lines are not linked to specific types of address information. The transformation treats each address uniquely and writes each address line to the first available port in this group.

- **Residue.** Use for unrecognized elements in the input address.

Match Code

Output port that summarizes the quality of the match between the processed address data and the address reference data.

Match Code Usage

Select Match Code to view a summary of the postal quality of the output address elements.

Use Match Code with the Element Result Status port. Element Result Status provides detailed information about the validity of each output port.

Port Location

The following table provides the location and default precision of the Match Code port:

Port Type	Port Group	Template Model	Precision
Output	Status Info	Basic	4

Match Code Output Codes

The following table describes the output codes generated on the Match Code port:

Code	Description
V4	The input data is correct. Address validation checked all postally relevant elements, and inputs matched perfectly.
V3	The input data is correct, but some or all elements were standardized. The input may contain outdated name data.
V2	The input data is correct, but some elements could not be verified because of incomplete reference data.
V1	The input data is correct, but user standardization has negatively impacted deliverability.
C4	Corrected. All postally relevant elements have been checked.
C3	Corrected. Some elements could not be checked.
C2	Corrected, but the delivery status is unclear due to absent reference data.
C1	Corrected, but the delivery status is unclear because user standardization has introduced errors.
I4	Data could not be corrected completely, but is very likely to be deliverable. There is a single match with an address in the reference data.
I3	Data could not be corrected completely, but is very likely to be deliverable. There are multiple matches with addresses in the reference data.
I2	Data could not be corrected. The address may be deliverable.
I1	Data could not be corrected and the address is not likely to be delivered.
Q3	Fast Completion Status. Address suggestions are available.
Q2	Fast Completion Status. Suggested address is complete but is mixed with

Code	Description
	elements from the input address.
Q1	Fast Completion Status. Suggested address is not complete. Add more inputs to the transformation.
Q0	Fast Completion Status. Insufficient input available to generate suggestions.
RB	Country recognized from abbreviation. Recognizes ISO two-character and ISO three-character country codes, and may also recognize common abbreviations such as "GER" for Germany.
RA	Country recognized from the Force Country setting in the transformation.
R9	Country recognized from the Default Country setting in the transformation.
R8	Country recognized from name without errors.
R7	Country recognized from name with errors.
R6	Country recognized from territory.
R5	Country recognized from province.
R4	Country recognized from major town.
R3	Country recognized from format.
R2	Country recognized from script.
R1	Country not recognized. Multiple candidate matches.
R0	Country not recognized.
S4	Parsing mode. The address was parsed perfectly.
S3	Parsing mode. The address was parsed with multiple results.
S2	Parsing mode. The address was parsed with errors. Elements changed position.
S1	Parsing mode. There was a parsing error. Input format mismatch.
N1	Validation Error. Validation did not occur because the country was not recognized.
N2	Validation Error. Validation did not occur because the required reference database is not available.
N3	Validation Error. Validation did not occur because the country data could not be unlocked.
N4	Validation Error. Validation did not occur because the reference database is corrupted or in the wrong format.
N5	Validation Error. Validation did not occur because the reference database is out of date.

Mailability Score

Output port that contains a single digit that represents the likelihood of successful delivery to the validated address, based on overall validation results.

Mailability Score Usage

Select Mailability Score to review code values that summarize the deliverability of the output addresses. The return codes are in a range from 0 through 5.

A return code of 0 represents an undeliverable address, and 5 represents an address that is certain of delivery.

Use the Match Code port with the Mailability Score port. The Mailability Score output codes provide additional information about addresses with Match Code values in the range I1 through I4.

Port Location

The following table provides the location and default precision of the Mailability Score port:

Port Type	Port Group	Template Model	Precision
Output	Status Info	Basic	2

Mailability Score Output Codes

The following table describes the Mailability Score output codes:

Code	Description
5	Fully certain of delivery
4	Almost certain of delivery
3	Confident of delivery
2	Reasonably confident of delivery
1	Delivery is doubtful
0	Undeliverable

Result Percentage

Output port that indicates the degree of similarity between an input address and an address in the reference data. The port represents the similarity as a percentage value.

Result Percentage Usage

Select Result Percentage to review the levels of similarity between the input addresses and the address reference data set.

The values range from 0 through 100. A value of 100 indicates that the address is a complete single match with an address in the reference data.

Port Location

The following table provides the location and default precision of the Result Percentage port:

Port Type	Port Group	Template Model	Precision
Output	Status Info	Basic	6

Result Percentage Output Codes

The Result Percentage output values range from 0 through 100. A value of 100 indicates that the address is a complete single match with an address in the reference data.